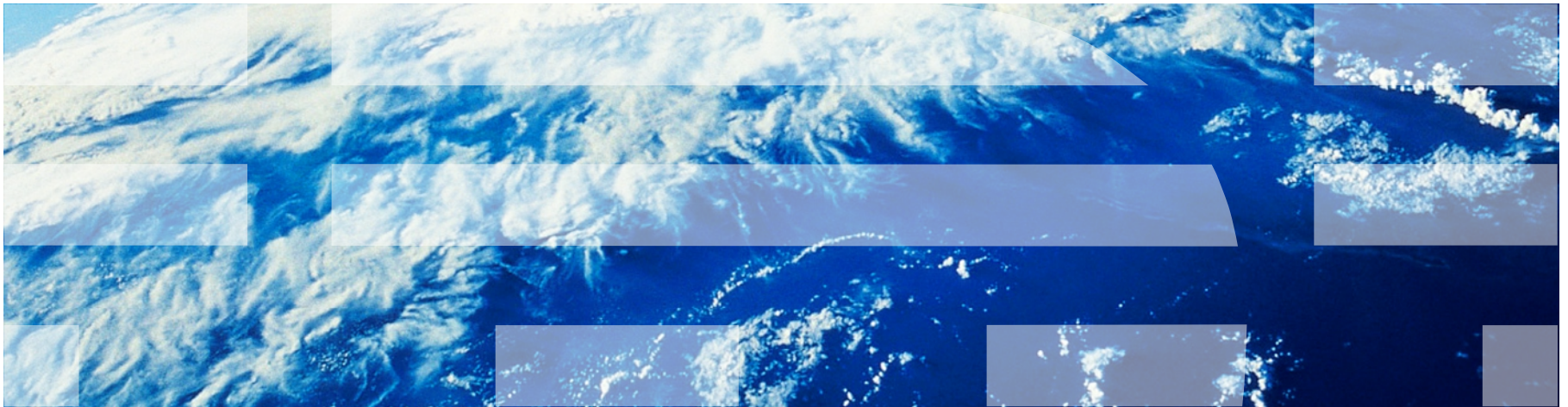
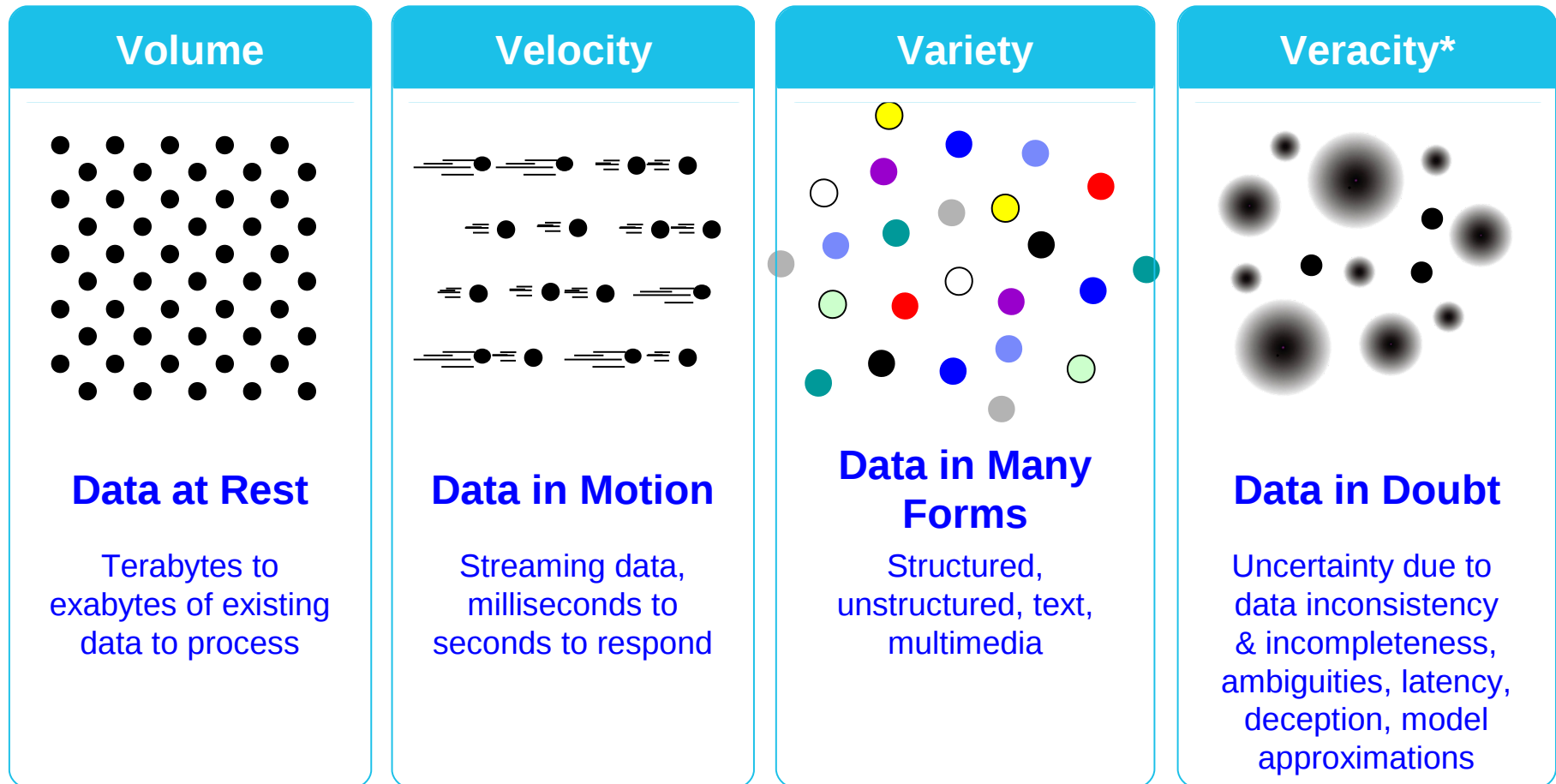


The Future of Big Data & Analytics:

Excerpts from IBM's Global Technology Outlook (2012 - 2013)



The fourth dimension of Big Data: Veracity – handling data in doubt

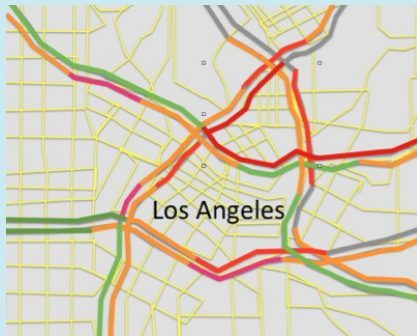


* Truthfulness, accuracy or precision, correctness

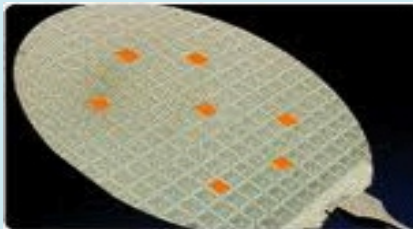
Uncertainty arises from many sources

Process Uncertainty

Processes contain
“randomness”



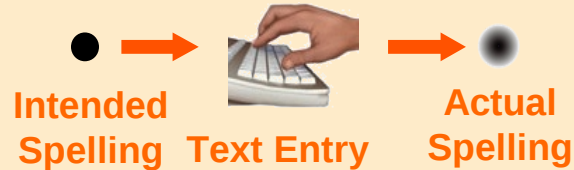
Uncertain travel times



Semiconductor yield

Data Uncertainty

Data input is uncertain



GPS Uncertainty



Testimony

{Paris Airport}

Ambiguity



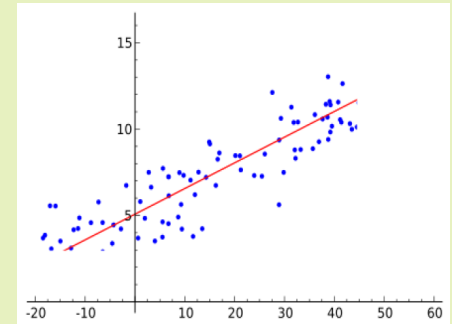
Contaminated?
Rumors

{John Smith, Dallas}
{John Smith, Kansas}

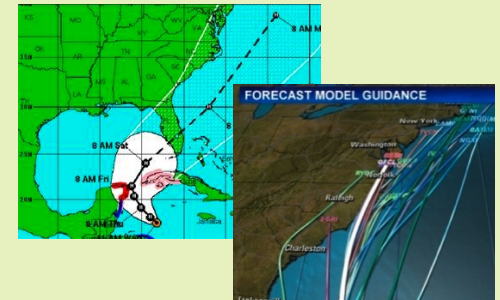
Conflicting Data

Model Uncertainty

All modeling is approximate

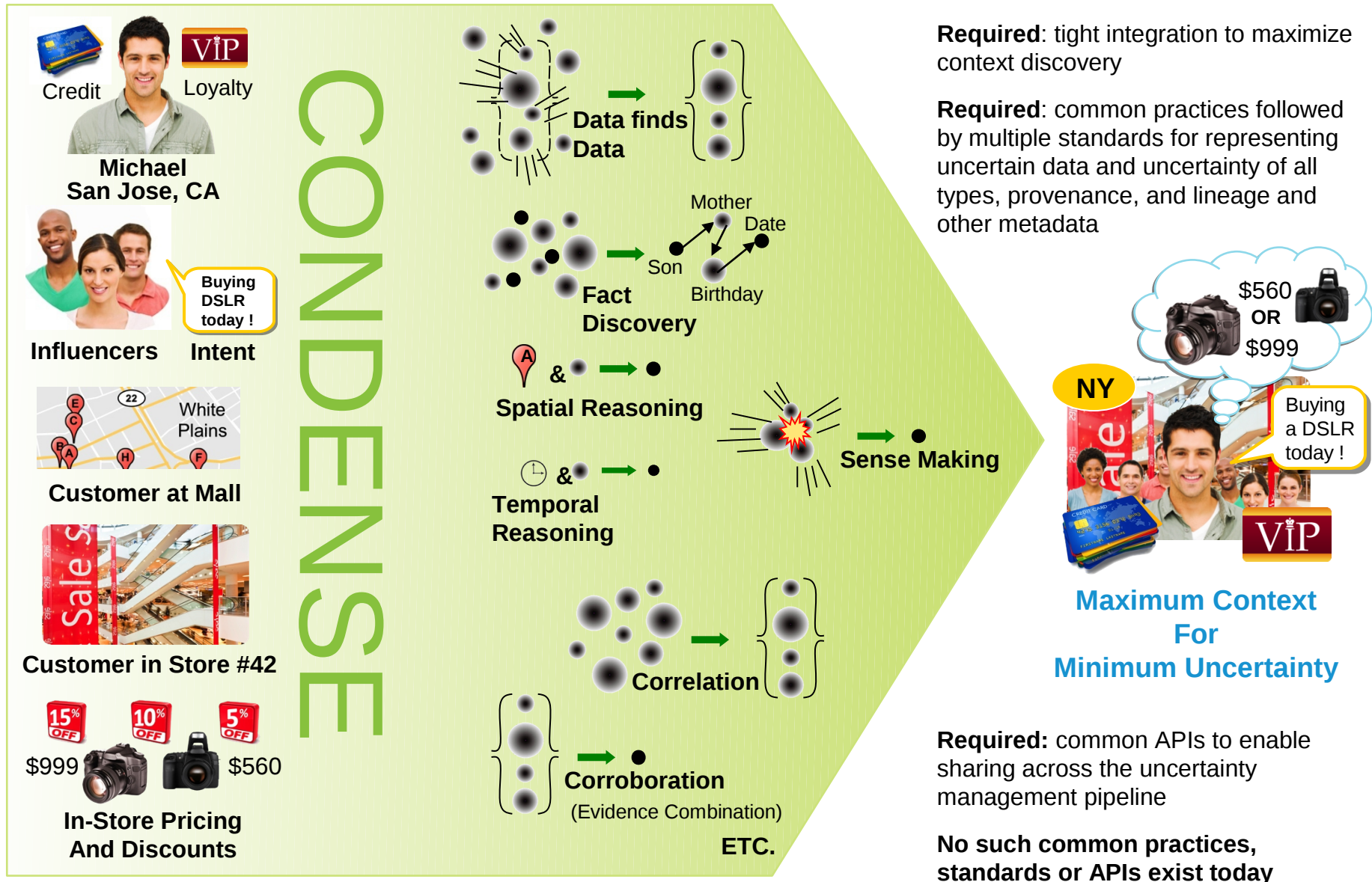


Fitting a curve to data



Forecasting a hurricane
(www.noaa.gov)

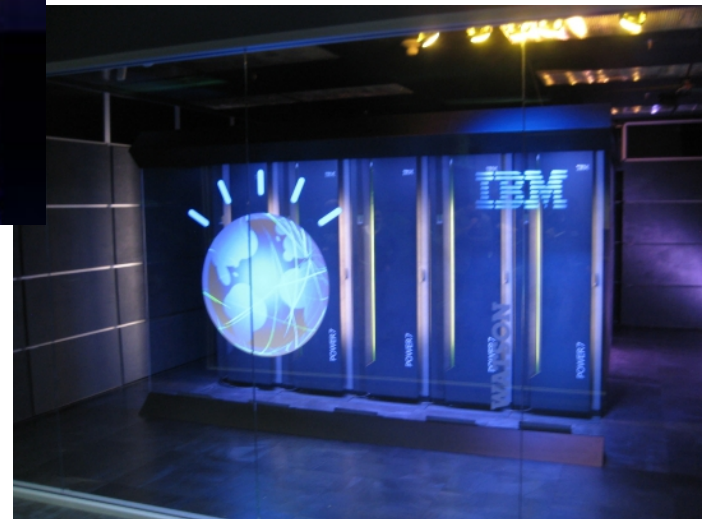
Condensing data reduces uncertainty by constructing context



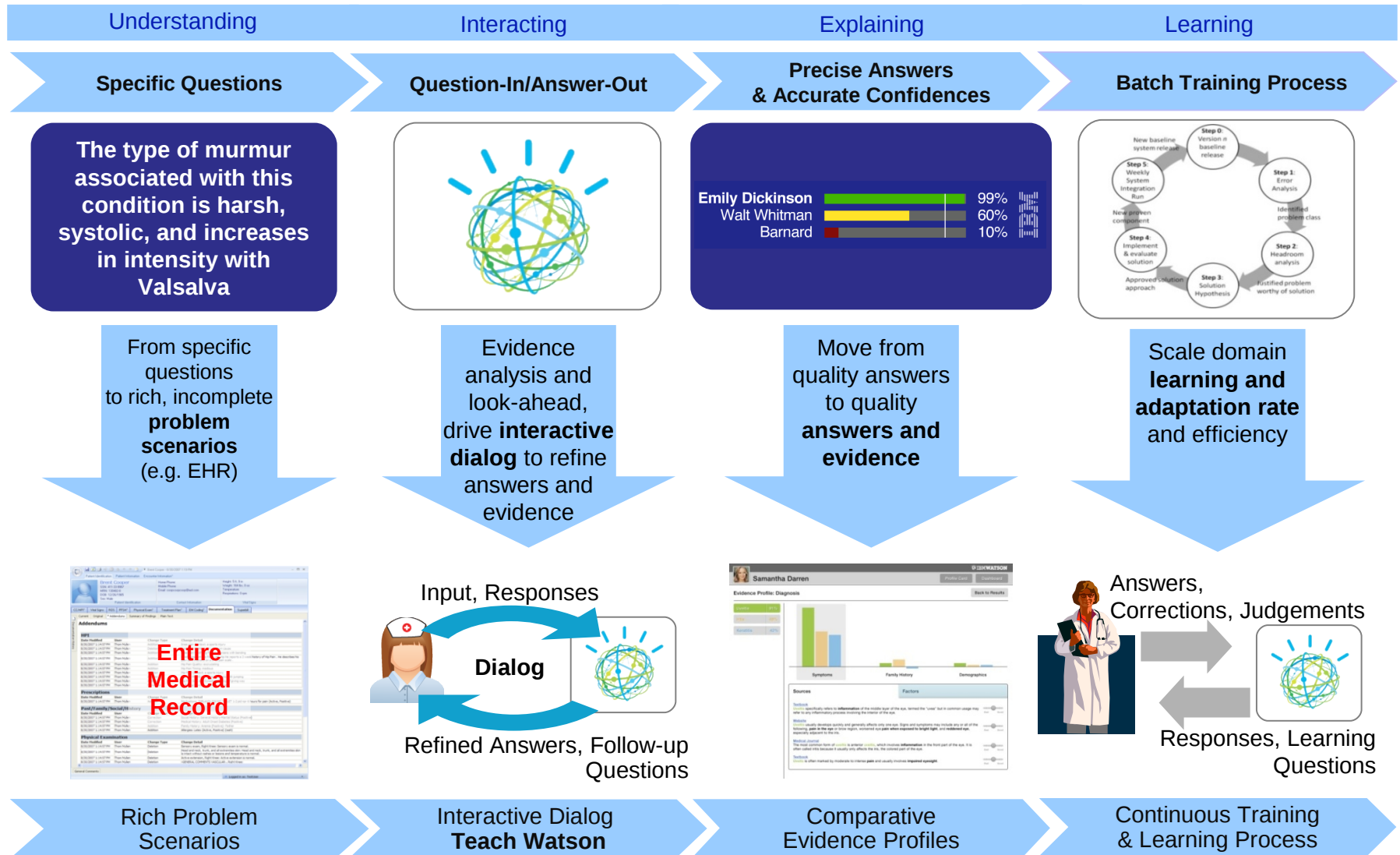
An IBM Grand Challenge



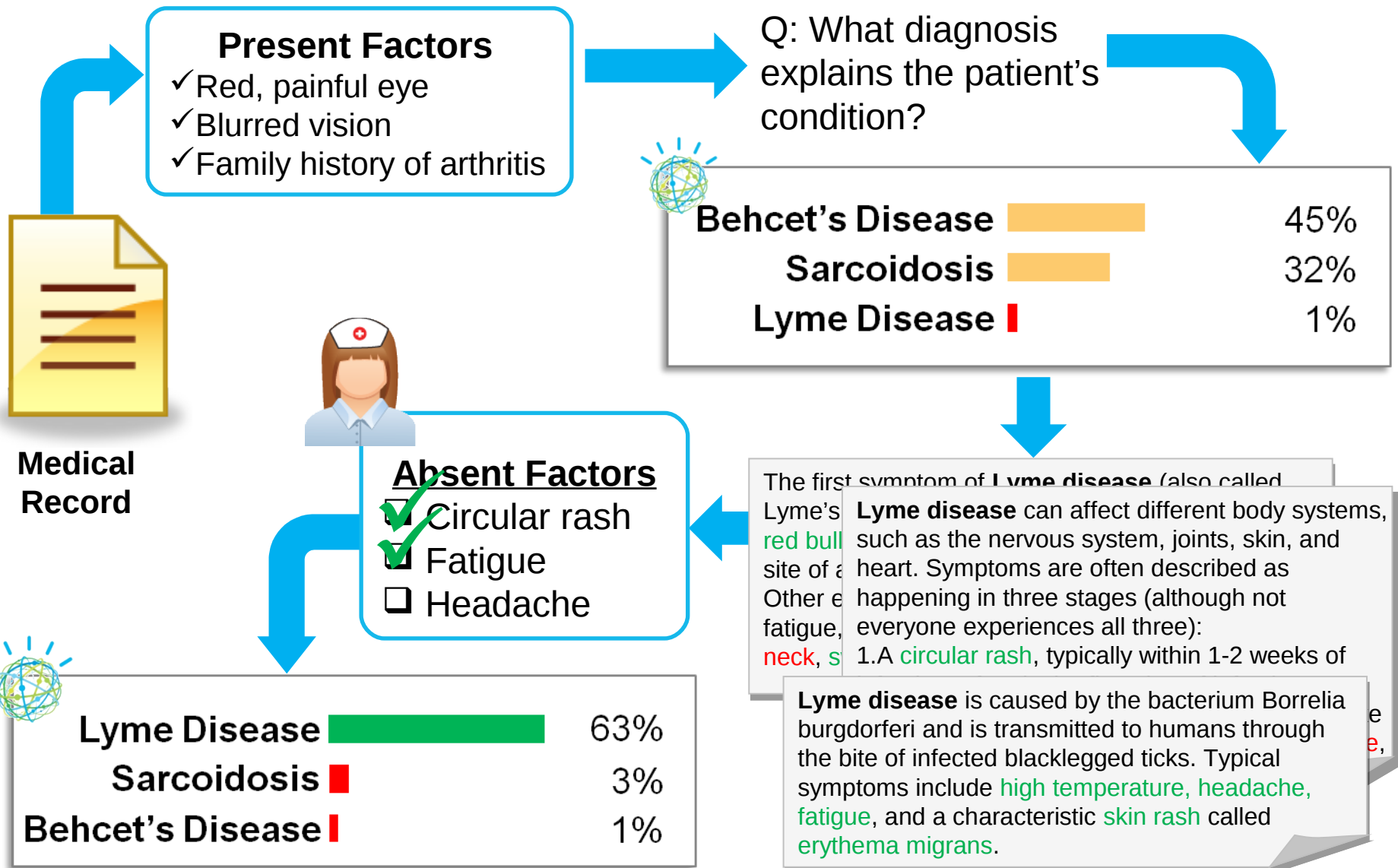
Build a system that rivals a human's ability to answer questions posed in natural language with accuracy, confidence, and speed.



Taking Watson beyond Jeopardy!



Dialoguing to an answer

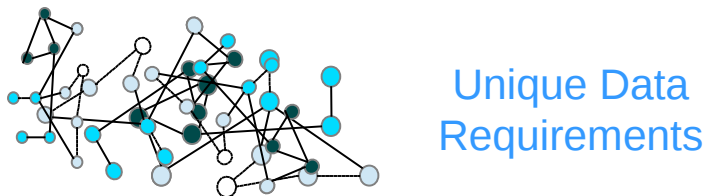


Context at Scale brings technical challenges

- Fully contextualized information will require at least 10x the storage of raw data.

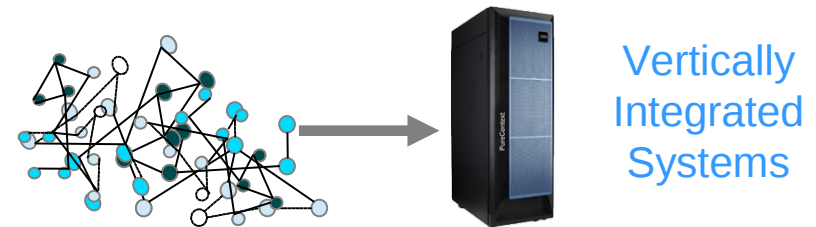


- Continual ingestion and curation will require continual deep analytics to discover new insights.
- Dynamic schema requirements and temporality will drive new database requirements.

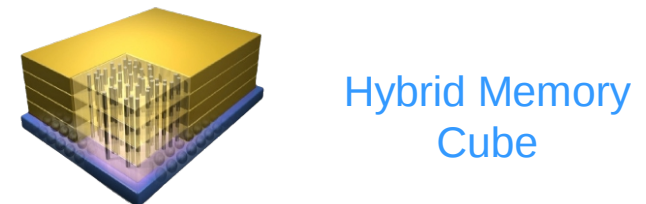


- High dimensional data with dynamic content and relations lead to irregular graphs that are notoriously hard to partition and favor large in-memory systems.

- Context never rests but data grows continuously with streaming high velocity input. Highly dynamic graphs will require ultra scalable data-structures, that support local graph traversal and at the same time representation/querying of global properties.



- Context size, dynamics, and access patterns will require data-centric, scale-in, highly integrated systems.



- Hybrid memory cubes and other disruptive technologies will enable large-scale, real-time, contextual processing.

